MODIS DATA STUDY TEAM PRESENTATION

March 9, 1990

AGENDA

1. Potential MODIS Data Loss: An Interview With Chris Scolese (McKay)
2. MODIS Data Packetization Considerations (Ardanuy)
3. Sea Surface Temperature (SST) (Wolford)
4. Update on Interdisciplinary Requirements for MODIS Data (Ardanuy)

Potential MODIS Data Loss:

An Interview with Chris Scolese

To develop information needed for the selection of an appropriate data packet structure for MODIS Level-0 data, Chris Scolese, EOS Systems Manager, was interviewed on the morning of Friday, March 2. This report first presents some background material related to the issues discussed and then the actual substance of the interview.

1. Background

Data loss mechanisms are potentially of decisive importance in determining the optimum MODIS Level-0 data packetization strategy.

In September of last year, the MODIS Data Study Team recommended a spectrally-oriented data packetization structure for MODIS Level-0 data. Each MODIS data packet would contain only data from a single spectral channel. This approach was recommended to facilitate selective routing of MODIS Level-0 data. It is thought that investigators or other near-real time users of Level-0 data will usually want data for some, but not all, spectral bands. Data packetized by spectral bands can be selected using information in the packet header. Other data packetization schemes (e.g., band interleaved) could require processing of the entire data stream returned from the MODIS instruments to obtain a reduced data stream containing only data for certain specific spectral bands.

Since the initial recommendation was prepared, analysis has shown that, if MODIS data is packetized using the recommended structure, and if MODIS data is randomly lost as whole packets, then the effect of these packet losses is increased by a factor of **n** [assuming packet losses are small and comparing with corresponding losses using band-interleaved packetization], where **n** is the number of concurrent spectral bands required to complete processing for a given product. Increased data product loss occurs only if a whole packet of MODIS data is lost in a single random event.

2. Substance of the Interview

The topic of the interview was data loss mechanisms affecting MODIS Level-0 data. The question addressed during the interview is:

"Do such random data loss events actually occur"?

It appears that the answer is **"No"**, i.e. the random loss of whole MODIS data packets is considered extremely improbable for the TDRSS/CDOS data link. CDOS will normally deliver blocks of data uncorrectable by Reed-Solomon coding in the "as-received" condition. Including such data, the expected bit error rate is 10^{-8} or less (less if current transmitter power specifications are retained). Random loss of bits or blocks of data is not expected, i.e. processing circuitry will nearly always generate some sort of (right or wrong) output. Any random data losses that do occur will be included in determinations of the overall bit error rate as specified above.

Data anomalies will normally occur as bit errors randomly distributed throughout the data stream. The only credible mechanism that could cause the random loss of whole packets of MODIS data would be a bit error affecting the address portion of the packet. This data could be misrouted. However, redundant checks of packet attributes other than address (e.g. length) will also be applied, so that whole-packet data loss even with an address error is considered improbable.

Direct broadcast of MODIS and other EOS data was a secondary subject of discussion. Two modes are envisioned and will probably be implemented: an "omni" antenna mode using an omni-directional antenna at both the transmitter and receiver and handling 10-15 Mb/s, and a high-gain antenna mode requiring Landsat type antennas and transmitting about 100 Mb/s. Potential data loss mechanisms for these links have not been considered. Perhaps data loss for these links will not be critical since complete TDRSS-delivered data may be generally used for standard data product generation and will be available for (later) fill-in of any glitches in the direct broadcast data. It is thought that single-band packetization of MODIS data will facilitate data selection for direct broadcast.

MODIS DATA PACKETIZATION CONSIDERATIONS

Here we present some preliminary considerations on the advantages and disadvantages of two scenarios for MODIS data packetization: band-interleaved and spectrally sorted.

1. DATA LOSS MECHANISMS

Three general types of data loss mechanisms are possible:

- uncorrectable errors within data packets
- systematic packet losses over orbital segments
- random packet losses

Each loss mechanism implies a different form of impact when propagated through the MODIS processing system.

1.1 Uncorrectable Errors Within Data Packets

MODIS, and perhaps most EOS instrument data, will be subject to "Grade-2" service, which guarantees a minimum bit error rate of 10^{-8}. Over a day, the joint MODIS-N and MODIS-T data volume will be approximately 0.8×10^{12} bits. This data volume implies that an error will occur about once every ten seconds, and that there will be no more than

(1) $0.8 \times 10^{12} \times 10^{-8} = 0.8 \times 10^{4}$ bit errors per day

If we assume that the data packet lengths will be 8.8×10^{3} bits, and that each of the bits errors will be located in a separate data packet, then there will be

(2) $0.8 \times 10^{4} \times 8.8 \times 10^{3} = 7.0 \times 10^{7}$ suspicious bits

The fraction of uncontaminated to total MODIS data will then be

(3) $[0.8 \times 10^{12} - 7.0 \times 10^{7}]$ bits / 0.8×10^{12} bits = 0.99991

In other words, in the absence of other data loss mechanisms, the anticipated bit error rate for MODIS from the platform to the TDRS ground station will guarantee a data completeness of better than 99.99%.

1.2 Systematic Packet Losses Over Orbital Segments

Resource conflicts, such as for TDRSS with manned missions (i.e., the space station or the space shuttle), platform maintenance, or instrument calibration may all result in systematic losses of MODIS Earth-viewing data. These losses may extend over substantial portions of an orbit. Because these losses are systematic, there is little difference in the effect of this type of loss mechanism on either packetization concept.

1.3 Random Packet Losses

The MODIS instrument data will flow across many interfaces, including:

- instrument to platform data system
- interfaces within the platform data system
- platform data system to Tracking and Data Relay Satellite (TDRS)
- TDRS to TDRS ground terminal
- TDRS ground terminal to Data Interface Facility (DIF)
- DIF to Data Handling Center (DHC)
- DHC to EOSDIS

Across these interfaces, as well as within each element of the total data system, there may be opportunities for resource access conflicts or packet collisions which could result in lost MODIS data packets. We have been told by the EOS Systems Manager that there will be no "significant" losses of this kind.

2. END-TO-END DATA COMPLETENESS

To our knowledge, there has been no end-to-end simulation of the loss characteristics for EOS instrument data. We are concerned here about the implications of lost data packets on geophysical product determination. Table 1 compares the advantages and disadvantages of the two data packetization strategies being considered. The spectrally sorted data packetization method is preferable, particularly for special types of data handling, provided that there is no adverse impact on the generation of the standard geophysical data that are the ultimate archive product of the instrument. The mechanism for increased impact of this concept results from the fact that many of the MODIS algorithms depend on multiple bands of data. The loss of any of these bands then makes the retrieval of the parameter impossible. Typically, from two to ten wavelengths are involved. For the case of six wavelengths, the effect of a 10^{-8} bit error rate scales linearly to yield an effective completeness of 99.95%.

The effects of random packet losses will be identical. If random packet losses occur at a rate of 10^{-8} (implying that as many packets will be unusable due to bit errors as to random losses), then the effective completeness of the final geophysical products for spectrally sorted data will be about 99.9%.

3. UNCERTAINTIES AND CANDIDATE REQUIREMENTS

The MODIS Science team has unambiguously stated that "systematic losses of instrument data are unacceptable." The Science team has not levied a quantitative requirement for data completeness in the event of random losses. However, it may not be possible to define a defensible data completeness requirement, which will ultimately depend on geophysical product accuracy products at a given space/time scale, without a rigorous system simulation for that product. In any case, a realistic estimate of this requirement should be stated.

If an end completeness of 99.9% is required, then we may state that:

- Systematic losses of instrument data are unacceptable.
- A bit error rate of 10^{-8} is required, and must apply to the entire data system from instrument to EOSDIS.
- A random packet loss rate of 10^{-8} is required, and must apply to the entire data system from instrument to EOSDIS.
- Spectrally sorted data packets are acceptable

If an end completeness of no less than 99.95% is required, then we may state that:

- Systematic losses of instrument data are unacceptable.
- A bit error rate of 10^{-8} is required, and must apply to the entire data system from instrument to EOSDIS.
- If band-interleaved instrument data packets are formed, a random packet loss rate of 10^{-8} is required, and must apply to the entire data system from instrument to EOSDIS.
- If spectrally sorted instrument data packets are formed, a random packet loss rate of <10-9 is required, and must apply to the entire data system from instrument to EOSDIS.

Table 1. Trade-offs in spectrally-sorted versus band-interleaved data packetization for MODIS.

PACKETIZATION METHOD	ADVANTAGE	DISADVANTAGE
SPECTRALLY SORTED (one spectral band per packet)	Facilitates data handling, particularly for direct broadcast, quick-look, and near-real-time usage.	Loss of a packet will result in increased geophysical product losses compared to band interleaved.
BAND INTERLEAVED (multiple spectral bands per packet)	Loss of a data packet will impact only a few 1-km elements of a scan.	Every packet must be decommutated to obtain a full scan for any specific spectral band. Loss of a packet will result in the loss of all data products for that portion of the scan.

SEA SURFACE TEMPERATURES (SST)

This sizing estimate is based upon the current AVHRR multi-channel algorithm used by NOAA. The thermal radiation is corrected for the attenuation due to water vapor in the atmosphere.

There are significant differences between the method sized here and that used for AVHRR data:

1. It is assumed that cloud free ocean pixels are identified using the appropriate flags. Hence, the computation to do this is not included in this estimate.

2. It is assumed that the calculation will be done on a pixel by pixel basis rather than averaging over a small region.

3. For sizing purposes, it is assumed that the technique will work in coastal regions and over the entire scan.

The AVHRR algorithm attempts to obtain SSTs from uncontaminated, representative pixels. Consequently, there are range tests performed to determine it the pixel should receive further processing. These tests are:

1. Computed SST should lie between -2° and 35° Celsius and not differ from monthly climatologies by more than 7° Celsius.

2. Satellite zenith angle is evaluated (GT 40°) to determine if an additional correction is required for water vapor attenuation.

3. A separate nighttime and daytime temperature algorithm is used. The nighttime values must agree within one degree of each other.

4. The visible reflectance from each pixel is checked to determine if it below a preselected value to eliminate sun glint contamination.

5. Nighttime SSTs are not calculated when the satellite zenith angle is greater than 45°.

6. If the satellite zenith angle is greater than 53°, the sea surface daytime SST is not calculated.

The equations used are discussed below.

The equation used for correcting water vapor attenuation is:

$$Ts = Ti + C1(Ti - Tj) + C2$$

For satellite zenith angles greater than 40 degrees, an extra correction term is required of the form (sec Z - 1). A corrected brightness temperature is calculated based upon the temperature difference between two window channels ([illegible]μm and 12μm wavelengths).

The SST algorithm is divided into a nighttime algorithm and a daytime algorithm. The nighttime algorithm consists of three separate calculations which must agree to within 1 degree Celsius.

$$SST1 = a*T3.7 - b*T11 - c$$
$$SST2 = d*T11 - e*T12 - f$$
$$SST3 = g*T11 + h(T3.7 - T12) - i$$

where T3.7, T11, and T12 are the brightness temperatures and SST1, SST2, and SST3 are the dual window, split window , and triple window SSTs respectively. The brightness temperatures are obtained by solving Planck's equation for temperature. This takes about 25 steps.

The daytime algorithm consists of:

$$SST4 = g*T11 + h(T11 - T12) - f.$$

The calculated temperature must lie between -2.0 and +35.0 degrees Celsius and not differ from monthly climatology by more than 7.0 degrees. The reflectance from the pixel is checked to determine if it is below a preselected value. It is a function of bidirectional reflectance (solar angle, satellite zenith angle, and relative azimuth). It is assumed that the locations where sum glint is a problem are also identified by a previously calculated flag.

It is necessary to compare the observed sea surface temperature with climatological values. This will require recovering that value from a data set. The exact number of operations depends on the size of the data set and the efficiency of the search algorithm[1]. Recovering the climatological value may require more computer resources than the calculation.

Step	Number of Steps
Read earth located pixel	1
Read pixel land/ocean flag	1
If land go to (next pixel) else (continue)	2
Read pixel cloud flags	3
It cloudy go to (next pixel) else (continue)	2
Read satellite zenith angle	1
If > 45 degrees go to (next pixel) else continue	2
Read solar zenith angle	1
If > 90 degrees go to (nighttime algorithm) else (continue)	2
If LT 75 degrees go to (daytime algorithm) else (continue)	2
If > 75 degrees and LT 90 degrees go to (next step)	2
Read near IR reflectance (3.55 to 3.93 microns ?)	1
If LT 1 percent go to (nighttime algorithm) else (next pixel)	2

[1]The B-Tree algorithm will find a single entry in a one million entry database in five disk accesses. The sea surface data base could have approximately 100 million entries. It is estimated that recovering the climatology can be done with approximately 150 operations.

Convert radiance to brightness temp. (3 X 25) 75

Calculating Atmospheric Attenuation due to Water Vapor.
Read C1 1
Read C2 1
Compute Ts 5

Night Time Algorithm
Computing Three Nighttime Estimates of SST using these equations

$$SST1 = a*T3.7 - b*T11 - c$$
$$SST2 = d*T11 + e*T12 - c$$
$$SST3 = g*T11 + h(T3.7 - T12) - i$$

Read a, b, c, d, e, f, g, h, i 9
Compute SST1 4
Compute SST2 4
Compute SST3 6

Note: These values must agree within one degree Celsius.

If SST1 - SST2, < 1 go to (next step) else (next pixel) 2
If SST1 - SST3, < 1 go to (next step) else (next pixel) 2
If SST2 - SST3, < 1 go to (next step) else (next pixel) 2

Note: Values must be > -2.0 degrees Celsius and LT 35 degrees Celsius

If SST3 LT 35 degrees go to (next step) else (next pixel) 2
If SST3 LT -2 degrees go to (next pixel) else (continue) 2
Read climatological value of SST 1
If SST > abs 7 go to (next pixel) else (next step) 2
Write nighttime SST, its position, observation time to database 1

Daytime SST Algorithm

If satellite zenith angle > 53 degrees go to (next pixel) else (continue) 2

Compute SST4

$$SST4 = g*T11 + h(T3.7 - T12) - i$$

Compute SST4 6

If SST4 LT 35.0 degrees go to (next step) else (next pixel) 2
If SST4 LT - 2.0 degrees go to (next pixel) else (continue) 2

Read climatological temperature 150

If SST4 - climatology > abs 7 go to (next pixel) else (continue) 2

Write SST4, its earth location, observation time. 1

Total (per pixel) night 294

T̶otal (per pixel) day 269

With the assumption that there are 12,656 (= 1,582 * 8) pixel's per scan, 40% day, 50% cloud cover, and 70% ocean, the requirement for this product is

1.3 MFLOP/scan